ReMoDiffuse: Retrieval-Augmented Motion Diffusion Model

Year: 2023 | Citations: 238 | Authors: Mingyuan Zhang, Xinying Guo, Liang Pan, Zhongang Cai, Fangzhou Hong

---

Abstract

3D human motion generation is crucial for creative industry. Recent advances rely on generative models with domain knowledge for text-driven motion generation, leading to substantial progress in capturing common motions. However, the performance on more diverse motions remains unsatisfactory. In this work, we propose ReMoDiffuse, a diffusion-model-based motion generation framework that integrates a retrieval mechanism to refine the denoising process. ReMoDiffuse enhances the generalizability and diversity of text-driven motion generation with three key designs: 1) Hybrid Retrieval finds appropriate references from the database in terms of both semantic and kinematic similarities. 2) Semantic-Modulated Transformer selectively absorbs retrieval knowledge, adapting to the difference between retrieved samples and the target motion sequence. 3) Condition Mixture better utilizes the retrieval database during inference, overcoming the scale sensitivity in classifier-free guidance. Extensive experiments demonstrate that ReMoDiffuse outperforms state-of-the-art methods by balancing both text-motion consistency and motion quality, especially for more diverse motion generation. Project page: https://mingyuan-zhang.github.io/projects/ReMoDiffuse.html